High-Speed Interconnect IP for High-Performance Computing and Enterprise Applications

Cadence Design Systems

ABSTRACT

With Moore's law slowing down and data center requirements for bandwidth, compute and storage growing exponentially, there are significant advancements being made to server architecture beyond the x86 architecture. To enable these advancements and utilize their full potential, we need efficient "plumbing" between different components in a server and between servers in a data center. This requirement has fueled a flurry of innovation with new application optimized standards such as CCIX, GenZ, and updates to existing standards such as PCIe, DDR and Ethernet. In this presentation, we will cover the upcoming interconnect standards such as CCIX, GenZ and key updates to existing standards such as PCI Express, optimized for the data center. We review the advantages of using integrated and fully-verified design IP to reduce the cost and time to market for developing the next generation data center hardware solutions. We will share with you our advanced PCIe/CCIX solution implemented 7nm that will achieve the performance/watt/$ required for the next generation of data centers. By using integrated and fully verified design IP and subsystems, semiconductor and system companies can focus on truly differentiating their products rather than worrying about all the building blocks of their design.

High-Speed Interconnect IP for High-Performance Computing and Enterprise Applications

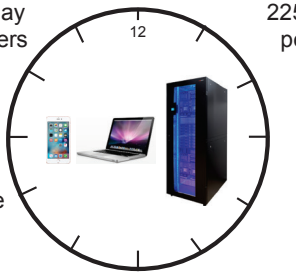
Sachin Dhingra, Marketing, PCI Express and Ethernet IP
TSMC OIP
Santa Clara, CA
September 13, 2017

cadence

Data Is Growing Exponentially

WeChat
40 minutes/day
1.1 Billion users

NETFLIX
225 GB/week
per person



YouTube
100 hours of video
uploaded each minute

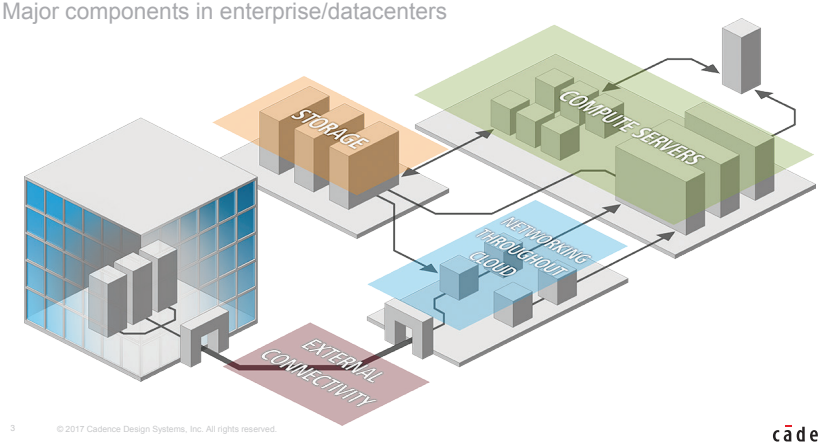
Amazon
426 items
sold/second

Facebook
4.5B posts/day
1B videos/day

cadence

Cloud and Infrastructure Market Vertical

Major components in enterprise/datacenters



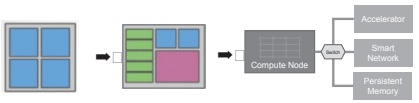
cadence

Implications of Data Growth

- Value of the insights gained from the data declines quickly
  - Need for High-Performance Compute (HPC)
- Innovation in HPC
  1. Increased performance (Moore's law)
  2. Multi-core architecture (parallel processing)
  3. Heterogeneous architecture
  4. Heterogeneous System Architecture (HSA)

- Businesses demanding real-time insight
- Increasing amounts of data to be analyzed

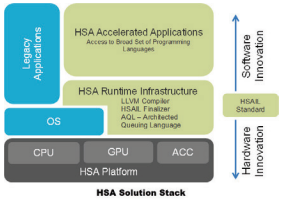
Ref: https://blog.dellemc.com/en-us/memory-centric-architecture-vision/



cadence

Heterogeneous Compute Architecture

- Combines dissimilar compute devices to achieve the best performance/$/watt
- Compute devices other than CPU include
  - Specialized CPUs
  - GPUs
  - FPGAs
  - DSPs
  - Network interface cards
  - Memory
  - Storage-class memory
  - Customer ASICs and other accelerators
- Cache/memory/storage hierarchy is rapidly becoming the bottleneck
  - 90% of the energy is wasted moving data



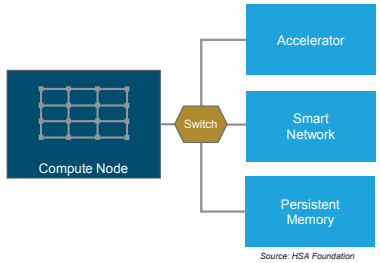
Source: HSA Foundation

cadence

Plumbing for Heterogeneous Compute Architecture

High-speed interconnect requirements

- High bandwidth
  - To open up the memory ↔ compute bottleneck
- Low latency
  - To meet the performance metrics for real-time analytics
- Flexibility
  - To scale up and scale out
- Easier programmability
  - Cache coherency
  - Simpler protocol
- Ease of deployment
  - Build on existing infrastructure such as PCIe or Ethernet



Source: HSA Foundation

cadence

cadence

High-Speed Interfaces Are “Plumbing” for HPC Platforms

- PCI Express (PCIe)
  - Primary interconnect on a server
  - Supports 16Gbps
  - Significant delays (3 yrs) going from PCIe 3.0 to PCIe 4.0
  - Just announced PCIe 5.0 supporting 32Gbps
- CCIX
  - Builds on PCIe designed for accelerators
  - Adds cache coherency
  - Increases bandwidth to 25Gbps and reduces latency
    - Looking at 32 and 56Gbps
- GenZ
  - Rack-level memory-semantic protocol (read-write)
  - Transport-layer protocol agnostic of PHY layer
  - Leverages PHY of Ethernet and PCIe

PCI EXPRESS

CCIX Cache Coherent Interconnect for Accelerators

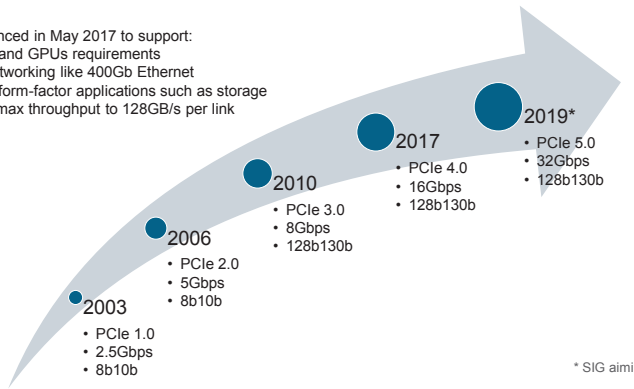
GEN Z

cadence

PCIe
Versatile, high-performance interconnect

PCIe 5.0 announced in May 2017 to support:
- Accelerator and GPUs requirements
- High-end networking like 400Gb Ethernet
- Constricted form-factor applications such as storage

It improves the max throughput to 128GB/s per link



* SIG aiming for 2018

cadence

PCIe - Markets and Applications
Solutions for a wide range of applications


Networking
- Network processors
- Ethernet host bus adapters


Compute
- HPC
- Desktops and laptops


Storage
- RAID
- SSD

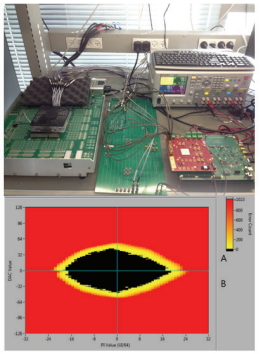

Industrial, medical, instrumentation, FPGAs
- Tester equipment
- FPGAs


Mobile
- WiFi (802.11an, 802.11ac, 802.11ad)
- Apps processor

cadence

PCIe 4.0 Interoperability: Cadence-Mellanox
PHY and controller

- 16Gbps PHY on TSMC 16nm process
- Electrical Interop – Mar'16
  - x4 r0.5
  - 16Gbps **16-Lane** T16FF+ PHY
  - Loss: 33dB
  - BER << 1e-15 with PRBS31
- Protocol Interop
  - x1 r0.5
  - Root Port: CDNS PHY+CTLR
  - End Point: Mellanox
- Passed 4.0 FYI compliance tests
  - Apr'17 and Aug'17

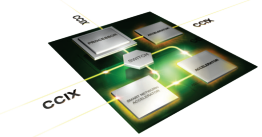

1. Electrical Interop


2. Protocol Interop

cadence

Cache Coherent Interconnect for Accelerators
CCIX

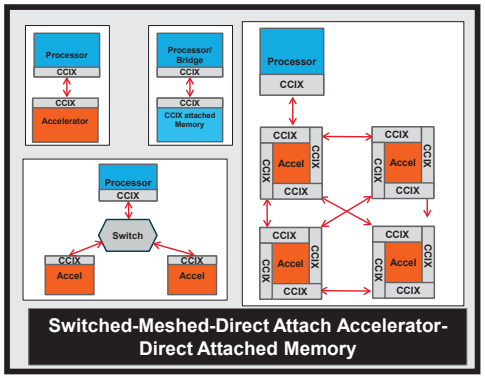
- Accelerates applications in the datacenter
  - Reduces power and cost, and improves efficiency
- Enables heterogeneous computing
  - Seamless data movement between devices
  - Cache coherency with devices with different ISAs
- Faster proliferation
  - Leverages existing PCIe ecosystem
- Higher bandwidth
  - CCIX defines 25GT/s (3X performance*)
  - Examining 56GT/s (7X performance*) and beyond
- Lower latency
  - Enabling low latency via light transaction layer
- Flexible and scalable
  - Point-to-point, daisy chained, and switched topologies
- Simpler programming
  - Driver-less, interrupt-less, and DMA-less framework for data sharing
- Open standard to promote innovation and reduce cost



cadence

CCIX: Applications and Topologies

- HPC
- In memory database processing
- Datacenter search
- Intelligent network acceleration
- Machine/deep learning
- 4G and 5G base-station
- Mobile edge computing
- Video analytics
- Embedded computing



cadence

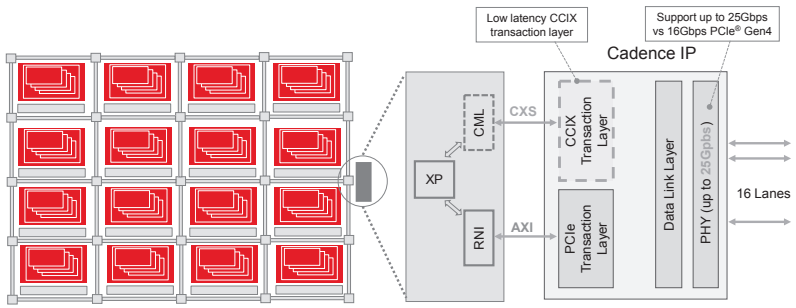
CCIX-25G Performance Today - Demonstration

- Transferring of a data pattern at 25 Gbps
  - Between two FPGAs
  - Across a channel comprised of PCIe CEM connector and a trace card
- Transceivers are electrically compliant with CCIX™
- Using the PCI Express infrastructure found in every data center
- Fastest data transfer between accelerators over PCI Express connections
- 3x faster transfer speed of PCI Express Gen3 solutions available today

Watch Demonstration at: https://youtu.be/JpUSAcnn7VA

cadence

Cadence CCIX Integration with a Key Partner



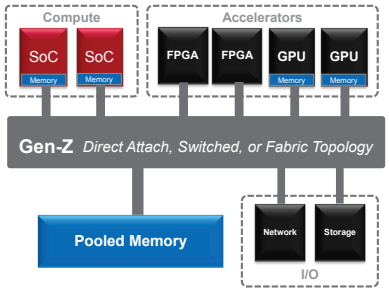
cadence

Gen-Z: A New Data Access Technology

High Bandwidth Low Latency
- Memory semantics – simple Reads and Writes
- From tens to several hundred GB/s of bandwidth
- Sub-100 ns load-to-use memory latency

Advanced Workloads and Technologies
- Real-time analytics
- Enables data-centric and hybrid computing
- Scalable memory pools for in-memory applications
- Abstracts memory media interface from SoC to unlock new media innovation

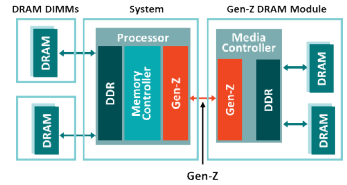
Secure Compatible Economical
- Provides end-to-end secure connectivity from node level to rack scale
- Supports unmodified OS / Apps for software compatibility
- Graduated implementation from simple, low cost to highly capable and robust
- Leverages high-volume IEEE and PCIe physical layers and broad, deep industry ecosystem
- Scales across multiple market segments

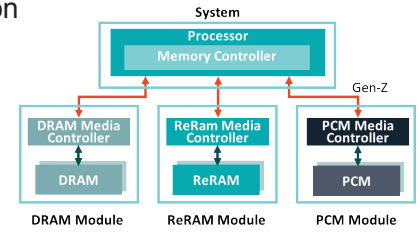


Co-Packaged-to-Enclosure-to-Rack Scale

cadence

Example Use Case: Memory Solution

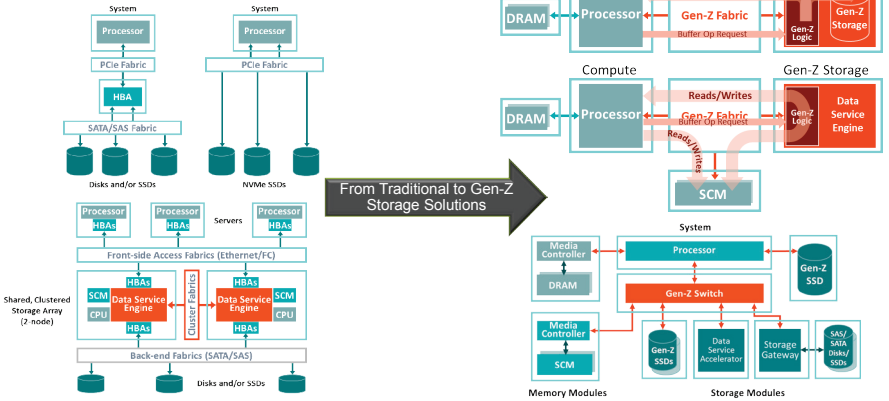




- Cost-effective, high-volume in-memory solutions
  - Legacy and new mechanical multi-TB modules
- Co-packaged, point-to-point, daisy-chain, or switch-based topologies
- Very high-bandwidth—up to 400+ GB/s modules
- Flattens memory / storage hierarchy
  - High-speed fabrics enable storage-free servers
- Multipath eliminates stranded memory, increases aggregate memory bandwidth, enables resiliency, etc.

cadence

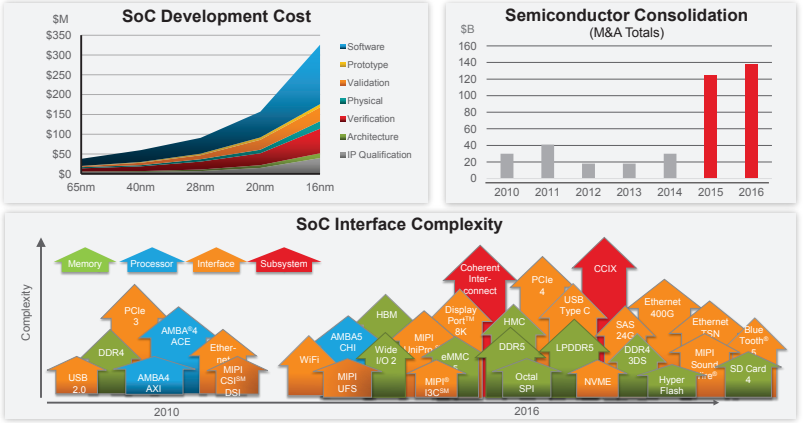
Example Use Case: Storage



cadence

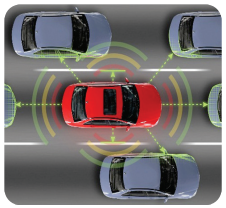
cadence

Forces Driving Commercial IP Adoption



cadence

Different Challenges for Different Markets

Datacenter	Mobile	Automotive
		
• Bandwidth • Latency • Reliability	• Power • Footprint • Performance	• Safety • Reliability • Bandwidth

cadence

Vertical Market Focus – Design IP

Alignment with TSMC process nodes

Datacenter

- **N7/N7+, 16FF+, 16FFC**
- DDR4@3200
- DDR5@4400
- HBM
- PCIe 4.0
- CCIX@25G
- Ethernet 10G-KR
- SATA3, USB3.1
- Flash Controllers

Mobile

- **N7, 16FFC, 12FFC, 28HPC+**
- LPDDR4/4X@4266
- LPDDR5
- MIPI D-PHY / M-PHY
- Flash
- USB3.1/3.0/2.0

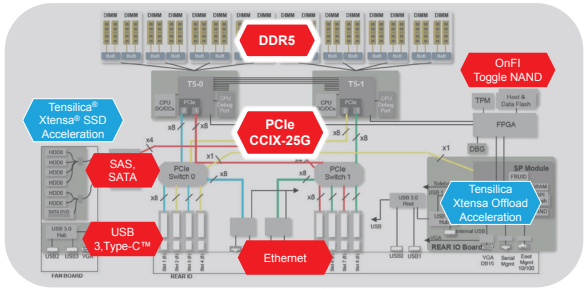
Automotive

- **N7, 16FFC**
- ISO 26262 ASIL-B ready
- TSMC9000A
- AEC-Q100
- Automotive Safety Features (ASF)

cadence

Cadence IP Datacenter Solution

- Earliest availability of high-bandwidth, low-latency protocols
  - PCIe 4.0, CCIX-25G DDR4 3200
- 16nm and 7nm process nodes
- Optimized for reliability, availability, and serviceability (RAS)



cadence

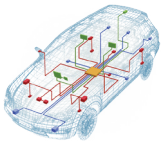
Recent Design IP Success

2 of 3 Top Mobile Phone Makers Use Cadence PCIe IP

1st Commercial PCIe Gen4 and CCIX in 7nm

DDR4 7nm Functional in Silicon

17 out of 25 Top Semiconductor Companies Use Cadence IP



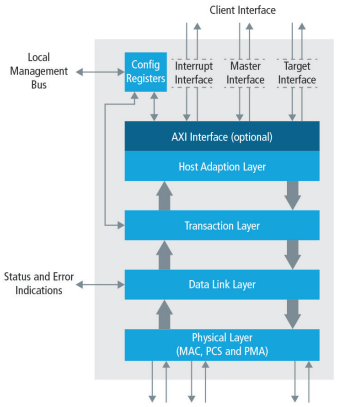
Automobile Leaders Adopting Portfolio in 7nm and 16FFC

cadence

Time to Market: Integrated PCIe Solution

Controller + PHY with drivers and VIP

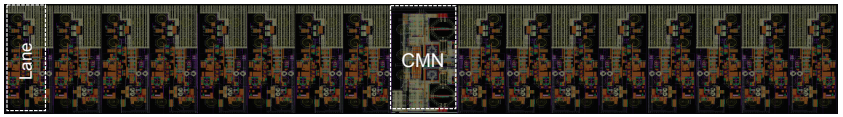
- Integration of controller and PHY
  - Time-consuming
  - Risky
  - Adds no value to the product
- Advantages of integrated solution
  - Pre-integrated and verified
  - Reduced risk and cost
  - Accelerated TTM
- Custom driver and documentation
  - Customized for each configuration



cadence

Cadence 25G Multi-Link/Multi-Protocol PHY IP (7nm)

Long reach (16G) / Medium reach (25G)



Highlights

- Test chip taped out
- Black-box delivery with lane margining and auto adaptation

Features

- Supports up to 16 lanes per macro at less than 2W
- PCIe 4.0 RX lane margining
- Long reach support with multi-tap DFE and auto-calibration
- Native bifurcation support per lane (x16 → 16 x1)
- Lower BOM cost with ref CLK and external resistor sharing
- Non-destructive on-chip oscilloscope
- Support for up to 2 reference clocks

Benefits

- **Performance:** Ultra-low TX jitter, high JTOL, and >30dB reach
- **Flexibility:** Multi-link enabled multi-protocol mix and match
- **Ease of use:** Pre-integrated PHY + controller solution

Standards

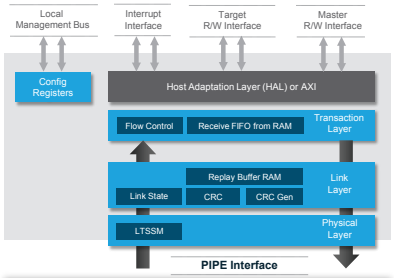
- CCIX up to 25Gbps
- PCIe 4.0 16Gbps with L1 sub-states and SRIS
- 25G-KR, 10G-KR
- USB 3.1, 3.0
- SATA 3
- XAUI/RXAUI, Q/SGMII
- And more…

cadence

Cadence Controller IP for PCIe

High performance, lowest power

PCI EXPRESS



IP Products

- Controller: Root ports, end points, and dual mode
- Software core driver and Linux reference driver
- Integrated solution: Controller, PHY, drivers, and VIP

Standards

- PCIe v1.1, 2.1, 3.1, and 4.0 (r0.5, 0.7)
- Intel PIPE v3.0 and 4.x
- SR-IOV v1.1

Benefits

- Complete PCIe solution: PHY, CTLR, VIP, and drivers
- Ease of use: Fully verified pre-integrated delivery
- Mature and silicon proven: Multiple designs in production
- Application optimized: High throughput and low-power ECN support
- Low risk: Fully verified and PCI-SIG compliant

Features

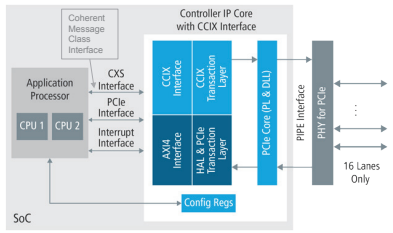
- Scalable design: 1 to 16 lanes and 256 PF/VFs
- Root port, Endpoint, and dual mode configurations
- High performance >95% link utilization
- Proven virtualization and bifurcation capabilities

cadence

Cadence Controller IP for CCIX

High performance, lowest power

CCIX Cache Coherent Interconnect for Accelerators



IP Products

- Controller: Root ports, end points, and dual mode
- Software core driver and Linux reference driver
- Integrated solution: Controller, PHY, drivers, and VIP

Standards

- CCIX r1.0 (16G)
- PCIe v1.1, 2.1, 3.1, and 4.0 (r0.5, 0.7)
- SR-IOV v1.1

Benefits

- Complete CCIX solution: PHY, CTLR, VIP, and drivers
- Ease of use: Fully verified pre-integrated delivery
- Silicon proven: Based on PCIe solution with multiple designs in production
- Application optimized: High throughput and low-power ECN support
- Low risk: Fully verified and interoperated with 3rd parties

Features

- 16 lanes with up to 256 PF/VFs
- High performance >95% link utilization
- 2 virtual channels: 1) AXI for PCIe 2) CXS for CCIX

cadence

Developing Solutions with TSMC

- The largest pure play foundry
  - Process and packaging technology leader
  - Wafer fab capacity leader in bleeding-edge nodes
  - Largest customer base
  - Broadest EDA and IP ecosystem
- TSMC enterprise leadership
  - Process technology: 7nm, 12nm, 16nm, 28nm
  - Packaging technology: CoWoS, InFO
  - Broad deployment of enterprise IP: DDR, PCIe, CCIX
- Cadence collaboration and partnership
  - Early enterprise IP engagement in advanced nodes
  - Silicon-proven flagship IP: DDR, PCIe, CCIX
  - Optimized tool flows
  - Collaboration on enterprise system solutions

29

cadence